685. On the performance of superposition window

Yanxue Wang, Jiawei Xiang, Zhanshi Jiang, Lianfa Yang
Guilin University of Electronic Technology
Guilin, 541004, China
E-mail: *yan.xue.wang@gmail.com*

(Received 31 October 2011; accepted 4 December 2011)

Abstract: Superposition window is often used in the digital signal processing and other fields of signal processing such as power spectral estimation and adaptive time-frequency analysis. Different overlap and windows used in superposition system may affect the final results. The main contribution of this paper is in providing the insight into the properties of the overlap-add technique with different window or overlap ratio, which is very helpful in selecting these parameters for a practical application.

Keywords: superposition frame, overlap-add, window functions.

1. Introduction

The individual signal is first windowed and split into some segments. The segmentation of a signal into non-overlapping segments can result in discontinuity artifacts at the segment edges. To reduce these artifacts, overlap-add technique can be applied to the windowed data. Overlapping technique is early used in the Welch's method to compute the modified periodogram method [1] and then expanded to synthesize a signal from its short-time Fourier transform [2] [3]. Modified overlap technique, such as weighted overlapped segment averaging method with proper overlap was applied to reduce the variance for the power spectral estimation [4]. A revisited Welch method via circular overlap is developed in the application of the nonparametric power spectrum power [5].

Most window functions afford more influence to the data at the center of the set than to data at the edges, which represent a loss of information. To mitigate that loss, the individual data sets are commonly overlapped in time or in frequency domain. Superposition window is adopted to solve the inconsistence problem S transform in time domain and frequency domain [6]. An adaptive, linear time-frequency analysis is proposed using superposition frame [7, 8]. An adaptive spectral kurtosis is developed by the author based on the superposition frame in frequency domain and is successfully used to detect the bearing fault [9]. Fixed overlap windows are employed in overlap-add procedures and a variable amount of overlap technique is proposed in [10].

Since overlap technique has been widely used, its performance of the different overlap and windows was seldom investigated. This paper gives the influence of the windows on the overlapping technique. The rest of this paper is organized as follows. Theory of the superposition frame is briefly introduced in section 2. Performance of the superposition frame is investigated in detail in section 3. Conclusions are given in section 4.

2. Theory of the superposition frame

Consider breaking an input signal x into frames using a finite, zero-phase, length M window w. then we may express the mth windowed data frame as

$$x_m(n) = x(n)T_{mR}w(n), n \in (-\infty,\infty) \tag{1}$$

where R is frame step (hop size), m is frame index and $T_{mR}w(\cdot)$ is the defined translation operator which can be written as

$$T_{mR}w(\cdot) = w(\cdot - mR) \tag{2}$$

The hop size is the number of samples between the begin-times of adjacent frames. Specifically, it is the number of samples by which we advance each successive window. Overlap ratio λ is determined by window length and hop size,

$$\lambda = 1 - \frac{R}{M} \tag{3}$$

when the overlap ratio is 50% that denotes $R = M/2$. For frame-by-frame spectral processing to work, we must be able to reconstruct x from the individual overlapping frames, ideally by simply summing them in their original time positions. This can be written as

$$x(n) = \sum_{m=-\infty}^{\infty} x_m(n) = \sum_{m=-\infty}^{\infty} x(n)T_{mR}w(n) = \frac{x(n)}{G(R)} \sum_{m=-\infty}^{\infty} T_{mR}w(n) \tag{4}$$

Hence, $x = \sum_m x_m$ if and only if,

$$\sum_{m\in\mathbb{Z}} T_{mR}w(n) = G(R), \forall n \in \mathrm{Z} \tag{5}$$

If $G(R)$ is a constant, this window w is a constant-overlap-add (COLA), which may keep perfect reconstruction in STFT. Via the Poisson summation formula, condition can be written as

$$\sum_{m\in\mathbb{Z}} T_{mR}w(n) = \frac{1}{R}\sum_{k=0}^{R-1} W(\frac{2\pi k}{R})e^{j\frac{2\pi kn}{R}} \tag{6}$$

Thus the COLA constraints in frequency domain [11]

$$W(\frac{2\pi k}{R}) = 0, \forall k \in \mathrm{Z} \cap k \neq 0 \tag{7}$$

In other words, a window w gives constant overlap-add at hop-size R if and only if the transformed window W is zero at all harmonics of the frame rate $2\pi/R$. Eq. (7) describes the weak COLA constraint in the frequency domain. When the short-time spectrum is being modified, these conditions no longer apply, and a stronger COLA constraint is preferable. The strong condition of COLA is to require that the transformed window $W(\omega)$ be band-limited consistent with down-sampling by R:

$$W(\omega) = 0, |\omega| \geq \frac{\pi}{R} \tag{8}$$

This condition is sufficient, but not necessary, for perfect COLA reconstruction. However, it cannot be achieved exactly by finite-duration window functions.

When either of the strong or weak COLA conditions is satisfied,

$$\sum_{m\in\mathbb{Z}} T_{mR}w(n) = \frac{W(0)}{R} \tag{9}$$

where $W(0) = \sum_{n=-\infty}^{+\infty} w(n)$. When Eq. (9) is satisfied, perfect reconstruction of the short-time Fourier transform is achieved, i.e., $y(n) = x(n)$, while $y(n)$ is written as

$$y(n) = \frac{R}{W(0)} \sum_{p=-\infty}^{\infty} \left[\frac{1}{N} \sum_{k=0}^{N-1} X_R(\frac{2\pi k}{N}) e^{\frac{j2\pi nk}{N}} \right] \tag{10}$$

3. Performance of the superposition frame

Parameters used to evaluate the performance of the superposition frame and the effect of the window type and overlap ratio are conducted in this section.

3.1. Parameters to evaluate the superposition frame

The most important features to measure the windows functions are mainlobe width, sidelobe attenuation, spectral leakage and gain.

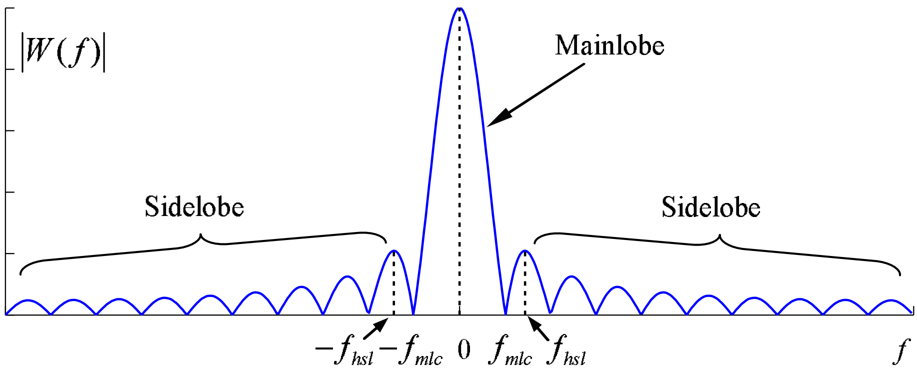

Fig. 1. Parameters of a window function

The mainlobe, sidelobe, cut-off frequency of the mainlobe, f_{mlc}, and the frequency associated with the height of the highest sidelobe, f_{hsl}, are shown in Fig. 1. To facilitate discussions, we denote $w(t)$ as a real, even and nonnegative temporal window and $W(f)$ as its Fourier transform. As is well known, $|W(f)|$ has a mainlobe at zero frequency and sidelobes on both sides (see Fig. 1).

3. 1. 1. Mainlobe width

The -3dB mainlobe width, denoted as $\mathrm{MLW}_{3\text{-dB}}$, is defined as the width of the mainlobe at 3dB below the mainlobe peak, and half of its value can be written as

$$0.5 \cdot \mathrm{MLW}_{3-\mathrm{dB}} \stackrel{\mathrm{def}}{=} \arg \max_{f} \{20 \log |W(f)/W(0)| \geq -3\} \tag{11}$$

$\mathrm{MLW}_{3\text{-dB}}$ mainly affects the frequency resolution in power spectral density estimation. Smaller $\mathrm{MLW}_{3\text{-dB}}$ leads to higher frequency resolution, and vice versa.

3. 1. 2. Relative sidelobe attenuation

The relative sidelobe attenuation (RSA) is defined as the log difference between the height of the mainlobe and the height of the highest sidelobe. It can be written as

$$\mathrm{RSA} \stackrel{\mathrm{def}}{=} 20 \log |W(f_{hsl})/W(0)| \tag{12}$$

The RSA denotes the difference between magnitude of the mainlobe and the maximum magnitude of the sidelobes. Undesirable spectral leakage can be reduced by increasing RSA.

3. 1. 3. Spectral leakage

Spectral leakage (SL) is defined as the ratio of the sidelobe power to the total window power and can mathematically be written as

$$\mathrm{SL} \overset{\text{def}}{=} 1 - \frac{\int_{-f_{mlc}}^{f_{mlc}} |W(f)|^2 \, df}{\int_{-\infty}^{+\infty} |W(f)|^2 \, df} \quad (13)$$

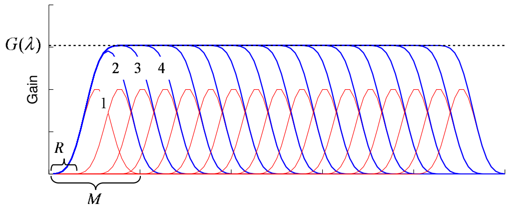

Fig. 2. Gain of the merged windows (Chebyshev window, λ=0.75). Red thin lines denote the initial window and its translation, while black thick lines are the merged windows. For interpretation of the references to the color in this figure legend, the reader is referred the web version of this article

3. 1. 4. Gain

The merged windows with initial width M and hop-size R are similar to a bandpass filter, as is shown in Fig. 2. Thus, the middle part is called as DC gain $G(\lambda)$ in this paper. According to Eq. (9), $G(\lambda)$ of the merged windows can be noted as

$$G(R) = \frac{W(0)}{R} \quad (14)$$

The thin lines in Fig. 2 are the shifted basic Hanning windows and thick lines denote the merged windows. Numbers shown in Fig. 2 refer to the merging times, for example, number 1 denotes the initial window itself and number 2 denotes two windows merge 1 times, and so forth. It can be observed that the gains of those windows, which merge at least two times, become a constant. This constant gain is necessary in keeping the perfect reconstruction of the signal processing. Actually, the merging times also affect the final gain of the merged windows which will be introduced in section 3. 4.

3. 2. The influence of the window functions

Window functions are often used in spectral analysis and so far many windows have been developed. Recently, Butterworth windows with two control parameters were proposed to achieve good performance in power spectral density estimation [12]. A performance comparison of window families can be found in [13]. Among them, Hanning, Hamming, rectangular, Bartlett, Kaiser, Gaussian and Chebyshev functions are frequently applied in practice. All the expressions of the mentioned window functions can be found in the Appendix. Fig. 3(a) shows the windows in time domain, while Fig. 3(b) gives the relating magnitude response. Superposition frame using different window has impact on mainlobe width, sidelobe attenuation, spectral leakage, and window overlap ratio. Such effects are discussed in this subsection, while the effect on the gain is given in next section 3.4. The effects of these 7 windows used in the superposition frame are investigated in this section.

Fig. 4(a) displays the effect of window merges based on different initial window functions.

As the number of merges increases, the $MLW_{3\text{-dB}}$ parameter of the merged window decreases. Similar pattern is observed for all the five window functions. It is also observed that the decreasing trends are bounded by that of the rectangle window function as is shown in the zoomed plot in Fig. 4(a).

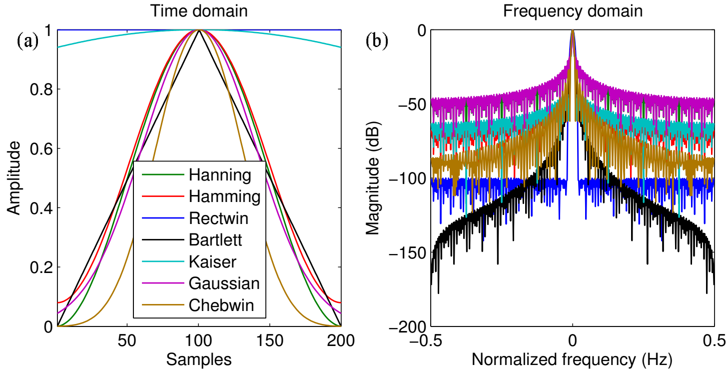

Fig. 3. Window functions and their magnitude responses, (a) time domain and (a, b) frequency domain (dB scale). Chebwin denotes Chebyshev window, Rectwin means Rectangular window (same as below). The color lines in (b) have the same meaning with those in (a). For interpretation of the references to the color in this figure legend, the reader is referred the web version of this article

The effect of window merges on the relative sidelobe attenuation is investigated using the same five window functions and is plotted in Fig. 4(b). It can be found that in Fig. 3(b) the RSA parameters of the merged Hanning, Hamming, Bartlett and Gaussian windows increase with the number of window merges. Once again, the trends are bounded by that of the rectangle window function.

As shown in Fig. 4(c), except the Chebyshev and rectangle window functions, the SL parameters for all the other five window functions increase with the number of window merging, moreover such increases are all bounded by the SL plot of the rectangle function.

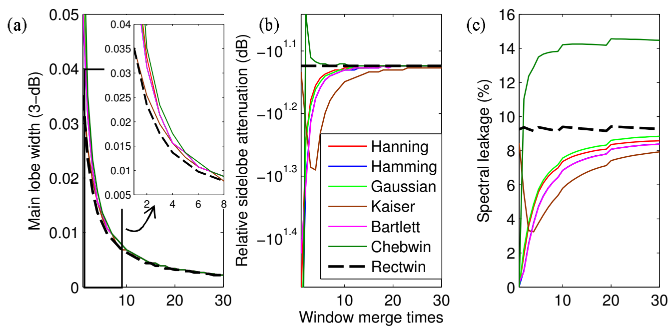

Fig. 4. (a) Main lobe width (3-dB), (b) relative sidelobe attenuation (dB) and (c) spectral leakage of the merged different type of windows. The color lines in (a) and (c) have the same meaning with those in (b). For interpretation of the references to the color in this figure legend, the reader is referred the web version of this article

3. 3. The influence of the variable overlap

In this subsection, the basic Hanning window is used as the initial window and different overlap ratios are evaluated in the superposition frame. We continue on to examine the effects of overlap ratio on $MLW_{3\text{-dB}}$, RSA and SL. Such effects (for Hanning window) are displayed in Fig. 5. It can be observed in Fig. 5 that when $\lambda \geq 0.5$ all the three parameters $MLW_{3\text{-dB}}$, RSA and SL show similar trends, either increasing (for parameters RSA and SL) or decreasing (for parameter $MLW_{3\text{-dB}}$) bounded by that of the rectangle window.

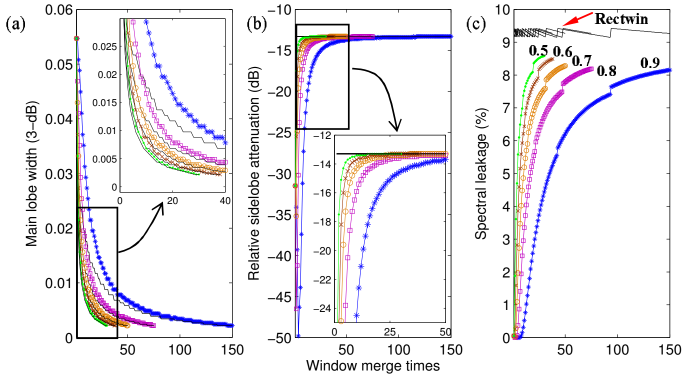

Fig. 5. (a) Main lobe width (3-dB), (b) relative sidelobe attenuation (dB) and (c) spectral leakage of the merged Hanning window and different overlap ratio. Thick black line denotes the rectangular window; green dark, brown, brown, purple and blue lines in (a), (b) and (c) denote λ=0.5, 0.6, 0.7, 0.8, 0.9, respectively. For interpretation of the references to the color in this figure legend, the reader is referred the web version of this article

3. 4. Gain of the superposition frame

As mentioned above, gain $G(R)$ of the folded windows can be computed using Eq. (14). That is to say, gain is related to the window function and hop size used in the superposition frame. Supposed the initial window length M is 20, hop size R changes from 1 to 20 (overlap ratio then reduces from 0.95 to zero according to the Eq. (3)). Fig. 6 displays the derived $G(R)$ with different windows and hop sizes. It can be determined that each $G(R)$ of different window decreases with the increase of R. When the overlap ratio λ is set to 0.5 (or $R = 10$), $G(R)$ functions of the merged windows are about 1 when using Hanning, Hamming, Bartlett, Chebyshev and Gaussian window as the initial window respectively, which is also called as a partition of unity [8, 11]. The accurate gains can be found in Table 1, where numbers in bold show they are near to 1 and N is the least superposition times needed.

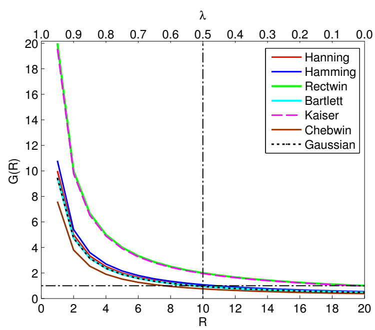

Fig. 6. Gain varies with hop size R in the process of merge when different window function is used. (Initial window length 20)

Table 1. Gain computed using different windows and overlap ratio (initial window width 20)

Windows	Hanning		Hamming		Rectwin		Bartlett		Kaiser		Chebwin		Gaussian	
λ	N	G(R)	N	G(R)	N	G(R)	N	G(R)	N	G(R)	N	G(R)	N	G(R)
0.75	3	2.10	3	2.06	4	4.00	3	1.89	6	3.91	3	1.52	4	1.89
0.50	**2**	**1.05***	**2**	**1.03***	2	2.00	**2**	**0.95***	2	1.96	2	0.76	**2**	**0.94***
0.25	2	0.69	2	0.68	2	1.33	2	0.63	2	1.30	2	0.50	2	0.63
0.00	1	0.50	1	0.52	**1**	**1.00***	1	0.47	**1**	**0.98***	1	0.38	1	0.47

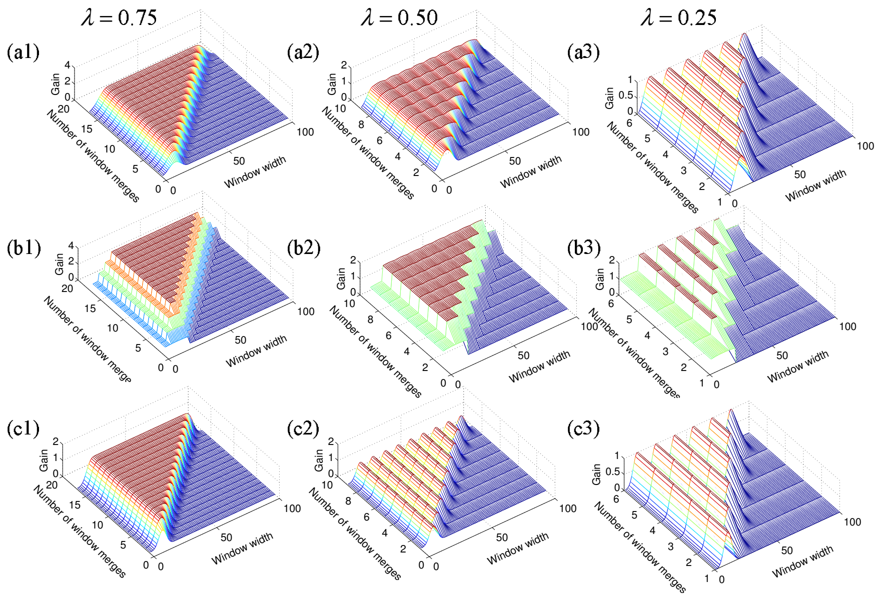

Fig. 7. Gain of merged different type of windows with three overlap ratio 0.75, 0.50 and 0.25: (a1) to (a3) using Hanning window; (b1) to (b3) using Kaiser window; (c1) to (c3) using Chebyshev window

Actually, the gains of the merged windows are not always constant (or called as COLA) if different ratio is used. This can be observed in Fig. 7, where Hanning, Kaiser and Chebyshev window functions are used in the superposition frame and overlap ratio is set to 0.75, 0.50 and 0.25, respectively. Gains of merged Hanning and Kaiser windows (shown in Fig. 7(a3) and (b3), respectively) are not constant in the case of $\lambda = 0.25$ and they change to constant when $\lambda = 0.50$ or $\lambda = 0.75$. However, gains of merged Chebyshev windows (shown in Fig. 7(c2) and (c3)) fluctuate both in the case of $\lambda = 0.25$ and $\lambda = 0.50$ and they are constant when λ is 0.75. To achieve the COLA of a predefined window type, there exists a boundary for the overlap ratio in the process of superposition. Fig. 8 displays the gains of merged Hanning windows when different overlap ratio is used (initial width is 20 samples, at least merging 5 times for each overlap ratio). Fig. 8 only gives the final merged windows whose width is 100 samples. It can be found that 0.5 may be considered as the boundary. When there is no limit of the merging times, numerical experimental results reveal the gains resulting from the merged windows such as Hanning, Hamming and Bartlett are constant (but gains are not necessarily equal to one) when overlap ratio is not less than 0.5 and the superposition times is large enough.

This COLA actually needs different superposition times for different overlap ratio. As the N values are shown in Table 1, those gains approximately to 1 (bold numbers of N values) are achieved using different merging times for the seven windows. Hanning window with width 20 is used as the initial window and λ changes from 0.5 to 0.95, $G(R)$ is shown in Fig. 9(a). To get the COLA gains (for Hanning window), the least superposition times for different λ values (thick line) are displayed in Fig. 9(b). It can be observed that Hanning window should at least merge three times when overlap ratio is 0.75. As overlap ratio increases, the merged times needed for COLA also increase.

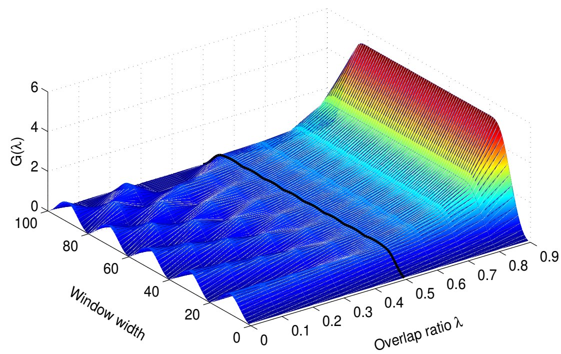

Fig. 8. Gains of the merged Hanning window with different overlap ratio. Black line denotes the overlap ratio 0.5

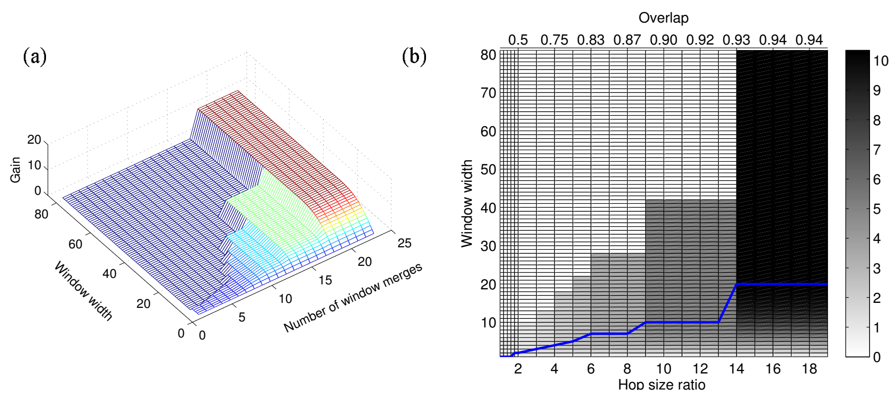

Fig. 9. Gain of the superposition Hanning window with different overlap ratio, (a) 3D show of the Gain, (b) 2D show of the Gain with different overlap ratio (log-scale); thick line is the marginal

4. Conclusions

Performance of the superposition frame depends on the chosen overlap ratio and the window function. Seven different window functions are investigated in COLA through numerical experiments in this paper. From the theory analysis and numerical experiment results, it can be concluded that:

(1) As the number of merges increases, the $MLW_{3\text{-dB}}$ of the merged windows decreases, while the RSA (except the Chebyshev and Kaiser) and SL of the merged windows increase. Besides, except for the Chebyshev window, three parameters of the other six window functions are all bounded by the rectangle window.

(2) The effects of different overlap ratio on $MLW_{3\text{-dB}}$, RSA and SL, such as Hanning window is

used as the initial window, parameters $MLW_{3\text{-dB}}$ will decrease in the process of superposition, while RSA and SL increase. All these parameters are also bounded by the rectangle window.
(3) COLA cannot be achieved exactly by finite-duration window functions, but it can be approximately implemented when the overlap ratio and superposition times are large, such as overlap ratio should be larger than 0.5 for Hanning window when this window merges at least 2 times.

Acknowledgments

This research is supported financially by the project of National Natural Science Foundation of China (No. 51105085, 51175097, 51165003).

References

[1] **P. Welch** The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, IEEE Transactions on Audio Electroacoustics 15 (1967) 70-73.
[2] **J. B. Allen, L. R. Rabiner** Unified approach to short-time Fourier analysis and synthesis, Proceedings of the IEEE 65 (11) (1977) 1558-1564.
[3] **L. D. Alsteris, K. K. Paliwal** Short-time phase spectrum in speech processing: A review and some experimental results, Digital Signal Processing 17 (3) (2007) 578-616.
[4] **G. C. Carter, A. H. Nuttall** On the weighted overlapped segment averaging method for power spectral estimation, Proceedings of the IEEE 68 (10) (1980) 1352-1353.
[5] **K. Barbe, R. Pintelon, J. Schoukens** Welch Method Revisited: Nonparametric Power Spectrum Estimation Via Circular Overlap, IEEE Transactions on Signal Processing 58 (2) (2010) 553-565.
[6] **S.-C. Pei, P.-W. Wang, J.-J. Ding, C.-C. Wen** Elimination of the discretization side-effect in the S transform using folded windows, Signal Processing, In Press, Accepted Manuscript (2011).
[7] **D. Rudoy, P. Basu, P. J. Wolfe** Superposition frames for adaptive time-frequency analysis and fast reconstruction, IEEE Transactions on Image Processing 58 (5) (2010) 2581-2596.
[8] **D. Rudoy, P. Basu, T. E. Quatieri, B. Dunn, P. J. Wolfe** Adaptive short-time analysis-synthesis for speech enhancement, presented at the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, 2008.
[9] **Y. Wang, M. Liang** An adaptive SK technique and its application for fault detection of rolling element bearings, Mechanical Systems and Signal Processing, Accepted (2010).
[10] **O. A. Niamut, R. Heusdens** Optimal time segmentation for overlap-add systems with variable amount of window overlap, IEEE Signal Processing Letters 12 (10) (2005) 665-668.
[11] **J. O. Smith** (2008). *Spectral audio signal processing*. Available: http://ccrma.stanford.edu/~jos/sasp/
[12] **T. H. Yoon, E. K. Joo** A Flexible Window Function for Spectral Analysis, IEEE Signal Processing Magazine 27 (2) (2010) 139-142.
[13] **N. Geckinli, D. Yavuz** Some novel windows and a concise tutorial comparison of window families, IEEE Transactions on Acoustics, Speech and Signal Processing 26 (6) (1978) 501-507.

Appendix

Equations of the windows mentioned in this work which can be also found in [13].
1) Rectangular window (Box-car)

$$w(n) = \begin{cases} 1, & 0 \le n \le N-1 \\ 0, & otherwise \end{cases} \quad \text{(A1)}$$

$$W(e^{j\omega}) = e^{-j\frac{N-1}{2}\omega} \frac{\sin\frac{N\omega}{2}}{\sin \omega/2} \tag{A2}$$

2) Bartlett window

$$w(n) = \begin{cases} \frac{2}{N-1}n, 0 \le n \le 0.5(N-1) \\ 2 - \frac{2n}{N-1}, 0.5(N-1) < n \le (N-1) \end{cases} \tag{A3}$$

$$W(e^{j\omega}) = \frac{2}{N}\left[\frac{\sin(N\omega/4)}{\sin(\omega/2)}\right]^2 e^{-j(\omega\frac{N+1}{2})} \tag{A4}$$

3) Hanning window (Raised-Cosine)

$$w(n) = 0.5\left[1 - \cos\left(\frac{2\pi n}{N-1}\right)\right], 0 \le n \le N-1 \tag{A5}$$

4) Hamming window

$$w(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) \tag{A6}$$

$$W(e^{j\omega}) = aW_R(e^{j\omega}) - \frac{1-a}{2}W_R(e^{j(\omega-\frac{2\pi}{N-1})}) - \frac{1-a}{2}W_R(e^{j(\omega+\frac{2\pi}{N-1})}) \tag{A7}$$

5) Kaiser window

$$w(n) = \frac{I_0\left(\pi\alpha\sqrt{1-\left(\frac{2n}{N-1}-1\right)^2}\right)}{I_0(\pi\alpha)} \tag{A8}$$

where $I_0(x) = 1 + \sum_{k=1}^{\infty}\left[\frac{1}{k!}\left(\frac{x}{2}\right)^k\right]^2$ is the zero-th order modified Bessel function.

6) Gaussian window

$$w(n) = e^{-0.5\left(\frac{n-(N-1)/2}{\sigma(N-1)/2}\right)^2}, \sigma \le 0.5 \tag{A9}$$

7) Chebyshev window (or Dolph-Chebyshev window, Dolph window) can be written in frequency domain as following

$$W(\omega_k) = \frac{\cos\left\{N\cos^{-1}\left[\beta\cos\left(\frac{\pi k}{N}\right)\right]\right\}}{\cosh\left[N\cosh^{-1}(\beta)\right]}, k = 0,1,2,3.\cdots, N-1 \tag{A10}$$

where $\beta = \cosh\left[\frac{1}{N}\cosh^{-1}\left(10^{\alpha}\right)\right]$.